The role of *cis*-regulatory elements in mutant mRNA of *FMR1* gene containing expanded CGG repeats in R-loop formation and regulation of noncanonical translation of pathogenic protein.

Daria Brygida Niewiadomska

ABSTRACT

The expansion of short tandem repeats located in either coding or non-coding regions of different genes underlies the pathogenesis of diverse human neurological diseases. The expansion of an unstable CGG repeat sequence within the 5' untranslated region (5'UTR) of *FMR1* has been implicated in the pathogenesis of multiple fragile X-linked syndromes.

Fragile X-associated tremor/ataxia syndrome (FXTAS) is a late-onset neurodegenerative disorder caused by the expansion of 55-200 CGG repeat, named as premutation of *FMR1*. The main symptoms of FXTAS are intention tremor, ataxia, and dementia. At the molecular level the disease is caused by the toxic *FMR1* mRNA that folds into a thermodynamically stable secondary structure at the region of excessively expanded CGG repeats (rCGGexp). Due to the sequestration by rCGGexp of many RNA-binding proteins the metabolism of RNA is highly disturbed and toxic inclusions containing this RNA are formed. The toxic mRNA with expanded CGG repeats is the template for non-canonical translation which results in the synthesis of toxic protein containing polyglycine tract (FMRpolyG) from the same mRNA from which the natural product of the *FMR1*, FMRP protein, is synthesized. Due to the strong aggregation properties, the protein is known to create intranuclear aggregates that lead to the disturbance of neurons and their death. Finally, the co-transcriptionally formed RNA:DNA hybrids, called R-loop structures, in the region of CGGexp are considered as another FXTAS pathomechanism leading to alterations in the transcription and driving DNA damage *via* cellular stress.

On the contrary, the fragile X syndrome (FXS) is associated with the expansion of more than 200 CGGs within the 5'UTR of the *FMR1* gene (name as full mutation) and is a neurodevelopmental disease, the most common form of inherited intellectual disability. The FXS patients are characterized by full mutation of *FMR1* which usually leads to the epigenetic silencing of the *FMR1* gene and consequently loss of FMRP protein. Although the silencing of *FMR1* is a complex process it has been shown that it is, at least partially, dependent on R-loops formation within CGGexp region.

The first part of the project aimed to establish the role of R-loops in the pathogenesis of FXTAS and FXS disorders. After confirming that R-loops are formed within 5'UTR of the *FMR1* in the

premutation range of CGG repeats the transcription efficiency regulated by the presence of R-loops was verified both *in vitro* and *in cellula*. Then, the contribution of short chemically modified antisense oligonucleotides (ASO-CCG) directly targeting CGGexp, involved in R-loop formation, on this structure stability and therefore on the *FMR1* transcription efficiency was tested. Finally, since *FMR1* full mutation in FXS leads to the silencing of *FMR1* transcription the long treatment with ASO-CCG was utilized to verify whether the reactivation of the *FMR1* transcription leading to the FMRP translation is possible. Results obtained in this part confirmed that R-loops forming within *FMR1* 5'UTR in FXTAS conditions have a negative effect on the *FMR1* transcription which can be partially abolished by ASO-CCG. However, according to FXS, ASO-CCG treatment did not reactivate the *FMR1* transcription from FXS cells which were characterized by full *FMR1* silencing. On the contrary, ASO-CCG treatment of FXS cells which possessed partially active *FMR1* locus resulted in the increased transcription rate of *FMR1* and its enhanced mRNA pool in the cytoplasm. Nevertheless, elevated *FMR1* mRNA level did not translate into increased FMRP level.

The second part of the project was concerning the *cis*-regulatory elements within *FMR1* 5'UTR and their involvement in the regulation of initiation of toxic FMRpolyG synthesis from near-cognate ACG or GUG start codons. In line with that, among others, the effect of different nucleotide sequence context in the vicinity of one of the near-cognate start codons on the FMRpolyG translation was established. Also, the effect of stable secondary RNA structure formed by the sequence located downstream of ACG near-cognate start codon on the translation initiation, and how different size of CGG repeats would affect the initiation of FMRpolyG synthesis were validated. Obtained data showed that both sequence context as well as stable secondary structure within mRNA have enormous effect on the initiation of FMRpolyG translation suggesting that this process is potentially regulated by many *cis*- and *trans*-factors targeting various regions/elements within the *FMR1* mRNA sequence.